Full-band quantum transport simulation in the presence of hole-phonon interactions using a mode-space *k·p* approach

OPEN ACCESS
IOP Publishing
Nanotechnology

Nanotechnology **32** (2020) 020001 (10pp)
https://doi.org/10.1088/1361-6528/abacf3

Full-band quantum transport simulation in the presence of hole-phonon interactions using a mode-space $k \cdot p$ approach

Hamilton Carrillo-Nuñez, Cristina Medina-Bailón, Vihar P Georgiev and Asen Asenov

School of Engineering, University of Glasgow, Glasgow, G12 8LT, United Kingdom

E-mail: hacarrillo@gmail.com

Received 5 June 2020, revised 15 July 2020
Accepted for publication 6 August 2020
Published 14 October 2020

Abstract

Fabrication techniques at the nanometer scale offer potential opportunities to access single-dopant features in nanoscale transistors. Here, we report full-band quantum transport simulations with hole-phonon interactions through a device consisting of two gates-all-around in series and a *p*-type Si nanowire channel with a single dopant within each gated region. For this purpose, we have developed and implemented a mode-space-based full-band quantum transport simulator with phonon scattering using the six-band $k \cdot p$ method. Based on the non-equilibrium Green's function formalism and self-consistent Born's approximation, an expression for the hole-phonon interaction self-energy within the mode-space representation is introduced.

Keywords: quantum transport, phonon scattering, non-equilibrium Green's function, nanowire transistors, six-band *k.p* Hamiltonian

(Some figures may appear in colour only in the online journal)

1. Introduction

The aggressive down-scaling of the complementary metal-silicon-oxide (CMOS) field-effect transistors (MOSFETs) has allowed major technological innovations, among them, novel materials [1], higher-κ gate dielectrics [2] and new structures such as gate-all-around (GAA) nanowire FETs (NW-FETs) [1, 3]. The latter is considered a very promising architecture, and a key component, for the next generation of transistors, both lateral and stacked GAA FETs [1, 3, 4]. According to the International Roadmap for Devices and Systems (IRDS) [4], the implementation of NW-FETs is a high priority for the industry. Hence, understanding their working principles/challenges, e.g. short-channel effects, is an urgent matter.

CMOS technology is also emerging as a promising asset for the realization of future quantum technologies, which combined with current nanometer fabrication techniques, can be used to fabricate multi-gate silicon structures with few dopants. A sketch of these devices is presented in figure 1, consisting of two GAA in series and a *p*-type Si nanowire channel with a single dopant within each gated region. Devices with similar architecture have shown promising applications in planar silicon-based spin qubits [5] or multi-dopant spectroscopy [6]. In order to provide insight into the working principles of these quantum devices, in this work, we study both coherence and dissipative quantum

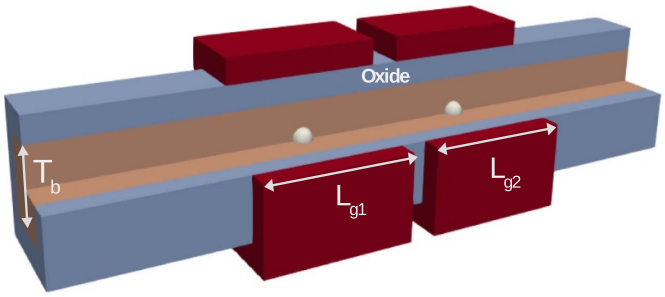

Figure 1. Simplified schematic of a silicon NW-FET nanowire with two gates. Distance between gates and the thickness of oxide layer is 1 nm. The latter is made from SiO_2. Gate lengths and silicon body thickness are $L_{g1} = L_{g2} = 7$ nm and $T_b = 3$ nm, respectively. Note that in each gated region there is a dopant, whose energy levels are independently controlled by the applied gate voltages.

transport phenomena through the device introduced in figure 1.

For accurate simulation of the electrical properties of NW-FETs, a full quantum treatment of the charge carriers is mandatory. A full quantum transport simulation involves the self-consistent solution of both Schrödinger and Poisson equations. Commonly, Schrödinger and its corresponding quantum kinetic transport equations are solved within the framework of the parabolic effective mass approximation (PEMA) of the band structure. However, the PEMA is most likely to fail to reproduce the valence band structure due to its high anisotropy. Quantum transport simulation of *p*-type nanodevices then requires a full-band description of the valence band structure.

Formalisms like tight-binding (TB) [7, 8], density functional theory (DFT) [9, 10], and k · p [11] are capable of computing the proper material band structure. While, the simulation of larger devices with atomistic quantum transport simulators becomes computationally costly, the $k \cdot p$ formalism, usually restricted to the Γ-point, offers a trade-off between accurate description of the band structure and much lower computational burden for device transport simulation. The most significant attribute of the $k \cdot p$ method is that it is based on six-bands for describing the valence band of Si and eight-bands for both the conduction and valence bands of direct band-gap semiconductors [12], e.g. III-V materials.

The $k \cdot p$ method has been used for simulating the transport properties of double-gate [13] and nanowire MOSFETs [14–16]. Comparison with TB simulations has shown that the two approaches are in agreement [17, 18]. However, only a few attempts at including phonon scattering within the $k \cdot p$ framework have been made [13, 18]. These studies, though, are based on a real-space representation of the $k \cdot p$ Hamiltonian, which is computationally costly.

In this work, we have developed a full-band quantum transport simulator based on the six-band $k \cdot p$ method that efficiently tackles hole-phonon interactions. This was achieved by combining the models proposed in [15] and [16]. Whereas, in the latter, the $k \cdot p$ Hamiltonian is projected into a much smaller subspace constructed by sampling the Bloch modes significant to the transport, the former allows a mode-space Hamiltonian via unitary transformation of the reduced-order Hamiltonian. The transport problem is then solved by means of the non-equilibrium Green's function (NEGF) technique and the hole-phonon interactions are included based on the self-consistent Born's approximation (SCBA) [19, 20]. Using the resulting mode-space Hamiltonian, a simplified expression for the hole-phonon interaction self-energy is introduced. This approach has been successfully implemented in the multi-scale transport simulation tool called Nano-Electronic Simulation Software (NESS) [21], which is currently under development at the University of Glasgow.

The paper is organized as follows. In section 2 and appendix A, the six-band $k \cdot p$ method, reduced-order and mode-space models are presented. Section 2 is also complemented by appendix B, where the elements of the $k \cdot p$ Hamiltonian discretized in K-space are listed. Section 3 briefly summarizes the main expressions of both NEGF and SCBA formalisms. Moreover, the phonon scattering self-energy within the mode-space representation is introduced. In section 4, we discuss our simulation results. Finally, conclusions are drawn in section 5.

2. Six-band $k \cdot p$ Hamiltonian

The three-band $k \cdot p$ Hamiltonian in $\{|X\rangle, |Y\rangle, |Z\rangle\}$ basis can be generally expressed as,

$$\begin{aligned} H'_{k\cdot p} =& Mk^2\mathbf{I} + (L-M)\left(Ak_x^2 + Bk_y^2 + Ck_z^2 + Dk_xk_y \right. \\ & \left. + Fk_yk_z + Gk_xk_z\right) - N\left(A'k_x^2 + B'k_y^2 + C'k_z^2 + D'k_xk_y \right. \\ & \left. + F'k_yk_z + G'k_xk_z\right), \end{aligned} \tag{1}$$

where $\mathbf{I}$ denotes the identity matrix and the $A, B, \ldots, G'$ matrices contain information about the crystallographic direction, as shown in appendix A. The Kane parameters, L, M, N, are given in terms of γ Luttinger parameters,

$$L = t_0\left(\gamma_1 + 4\gamma_2\right),\ M = t_0\left(\gamma_1 - 2\gamma_2\right),\ N = 6t_0\gamma_3, \tag{2}$$

where $t_0 = -\hbar^2/2m_0$ and m_0 is the free electron mass. Then, by choosing the spin-dependent basis $|\lambda S\rangle$ with $\lambda = X, Y, Z$ and $S = \uparrow, \downarrow$ for spin-up and spin-down spinors, the six-band Hamiltonian takes the form,

$$H_{k\cdot p} = \begin{pmatrix} H'_{k\cdot p} & 0 \\ 0 & H'_{k\cdot p} \end{pmatrix} + \begin{pmatrix} \mathbf{M} & \mathbf{N} \\ -\mathbf{N}^* & \mathbf{M}^* \end{pmatrix}. \tag{3}$$

The last term accounts for the spin–orbit interactions with,

$$\mathbf{M} = \frac{\Delta}{3}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{N} = \frac{\Delta}{3}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -i \\ -1 & i & 0 \end{pmatrix}, \tag{4}$$

where Δ is the spin–orbit splitting parameter.

Reduced-order and mode-space Hamiltonian

By applying an external potential $V(\mathbf{r})$ the periodicity of nanostructures is broken. The hole wavefunctions must then be obtained from the envelop equation,

$$H_R(-i\nabla,\mathbf{r})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) = [H_{k\cdot p}(-i\nabla) + V(\mathbf{r})\mathbf{I}]\,\Psi(\mathbf{r}), \quad (5)$$

where the total wavevector $\mathbf{k}$ has been replaced by its differential operator. The Hamiltonian H_R in equation (5) can be discretized in real space and implemented on NEGF solvers [13, 14]. Nevertheless, the real-space representation of H_R is computationally costly, particularly for the simulation of larger devices. We have adopted an alternative approach for the discretization of H_R by exploiting both the reduced-order [16] model and mode-space [15] representation. This is summarized as follows.

The Hamiltonian H_R is first transformed to K-space by the unitary transformation proposed in [15]. After a proper discretization in the transport direction, with N_x cross-sectional planes, the K-space Hamiltonian can be written as,

$$H_K(j) = h_0 + V_K(j). \quad (6)$$

The unperturbed Hamiltonian $h_0 = H_{0\,\mathrm{K}} + W_K + W_K^\dagger$, the coupling Hamiltonian W_K and potential $V_K(j)$ of the j-th cross-sectional plane are matrices of size N_bN_K. N_b is the number of bands considered in the $H_{k\cdot p}$ Hamiltonian and N_K is the number of mesh points in the K-space. N_K depends on the nanowire body thickness and the mesh space in the y- and z-directions, considering the x-direction as the transport direction. N_K can be optimized by constructing a reduced basis [16], as small as possible, able to accurately generate the original band structure within the energy range of interest, $E_n \lesssim E \leq E_\mathrm{v}$, from the top of the valence band edge E_v up to a cut-off energy E_n.

The reduced basis is constructed by solving the eigenvalue problem, coming from equation (6) when no external potential is applied,

$$\left(H_{0\,\mathrm{K}} + W_K e^{ik_x\Delta_x} + W_K^\dagger e^{-ik_x\Delta_x}\right)|\Psi_m\rangle = E|\Psi_m\rangle, \quad (7)$$

at different k_x sampled points. Filtering the eigen-energies lying inside the energy window of interest, a matrix with their corresponding eigenmodes Ψ_m can be constructed, being then **QR** factorized [16]. Using $\mathbf{Q}$ as a unitary transformation matrix, the reduced-order Hamiltonian discretized in K-space can then be obtained:

$$\tilde{H}_K(j) = \tilde{h}_0 + \mathbf{Q}^\dagger V_K(j)\mathbf{Q} = \tilde{h}_0 + \tilde{V}_K(j), \quad (8)$$

where $\tilde{h}_0 = \mathbf{Q}^\dagger h_0 \mathbf{Q}$. Note that the aforementioned approach becomes equivalent to the one proposed in [15] when only equation (7) is solved at $k_x = 0$.

The K-space reduced-order Hamiltonian might still be large for efficient full-quantum transport simulation. Instead, a mode-space representation of equation (8) can be employed, leading to,

$$\tilde{H}_M(j) = \tilde{H}_{0\,\mathrm{M}}(j) + \tilde{W}_M(j) + \tilde{W}_M^\dagger(j), \quad (9)$$

where,

$$\tilde{H}_{0\,\mathrm{M}}(j) = U_M^\dagger(j)\left(\tilde{H}_{0\,\mathrm{K}} + \tilde{V}_K(j)\right)U_M(j), \quad \tilde{W}_M(j) = U_M^\dagger(j)\tilde{W}_K U_M(j+1) \quad (10)$$

are $N_M \times N_M$ matrices and the unitary transformation matrix $U_M(j)$ consists of N_M mode wave functions $|\Psi_m(j)\rangle$ of size N_bN_K obtained from the reduced eigenvalue problem $\tilde{H}_K(j)|\Psi_m(j)\rangle = E|\Psi_m(j)\rangle$. The overall approach has been, therefore, transformed into finding the lowest $N_M \ll N_bN_K$ eigenmodes relevant for transport simulation. For the latter, equations (9) and (10) have been implemented into a mode-space NEGF solver, as follows.

3. Non-equilibrium Green's function approach with hole-phonon interactions

Within the mode-space framework presented in section 2, the retarded and lesser Green's functions for the active device region are written, in matrix notation, as,

$$G^R(E) = \left[E\mathbf{I} - \tilde{H}_M - \Sigma_C^R(E) - \Sigma_S^R(E)\right]^{-1}, \quad (11)$$

$$G^<(E) = G^R(E)\left[\Sigma_C^<(E) + \Sigma_S^<(E)\right]G^{\dagger R}(E), \quad (12)$$

respectively. The Hamiltonian $\tilde{H}_M$ and retarded/lesser self-energies $\Sigma^{R/<}$ are $N_xN_M \times N_xN_M$ matrices. $\Sigma_C^{R/<}$ is the self-energies for the semi-infinite leads, whereas, $\Sigma_S^{R/<}$ accounts for all scattering mechanisms. In the case of both acoustic and optical phonons, $\Sigma_S^{R/<} = \Sigma_{ac}^{R/<} + \Sigma_{opt}^{R/<}$. Hole-phonon interactions are tackled perturbatively within the SCBA [19, 20].

In practice, the retarded/lesser Green's functions are computed by exploiting a recursive algorithm that allows us to access their most relevant elements for the calculation of the physical quantities, such as the hole density,

$$p(x_j, y_m, z_n) = \frac{-i}{\Delta x_j}\int \frac{\mathrm{d}E}{2\pi}\mathrm{Tr}\left[p_R(x_j, y_m, z_n; E)\right], \quad (13)$$

where $\Delta x_j = x_{j+1} - x_j$ is the distance between adjacent planes. In equation (13), the following transformations,

$$p_R(x_j, \mathbf{r}_{nm}; E) = U_K(\mathbf{r}_{nm})p_K(x_j; E)U_K^\dagger(\mathbf{r}_{nm}), \quad (14)$$

$$p_K(x_j; E) = U_M(j)G_{j,j}^<(E)U_M^\dagger(j), \quad (15)$$

must be performed to recover the hole density from mode-space to real-space representation. $U_K(\mathbf{r}_{nm})$ is the unitary matrix [15] to transform from K-space to real-space and $\mathbf{r}_{nm} = (y_m, z_n)$ was used for brevity.

In the case of the current, on the other hand, no transformation is needed. It can be straightforwardly computed from,

$$I(x_j) = \frac{e}{\hbar}\int \frac{\mathrm{d}E}{2\pi}\mathrm{Tr}\left[\tilde{W}_M(j)G_{j+1,j}^<(E) - G_{j,j+1}^<(E)\tilde{W}_M^\dagger(j)\right]. \quad (16)$$

The Green's function elements $G_{j,j'}^<$ are $N_M \times N_M$ block matrices.

3.1. Hole-phonon interaction self-energy within the mode-space representation

Assuming bulk hole-phonon interactions, the retarded/lesser phonon self-energy in mode-space can be expressed as,

$$\Sigma^{R/<}_{p,p'}(x_j,E) = i\,\mathrm{Im}\left\{\sum_{n,n'} F^{n,p}_{n',p'}(x_j)\,\overline{G}^{R/<}_{n,n}(x_j,E)\right\}, \quad (17)$$

where n (n') and p (p') are mode indices. For acoustic phonons,

$$\overline{G}^{R/<}(x_j,E) = G^{R/<}_{j,j}(E), \quad (18)$$

whereas,

$$\overline{G}^{<}(x_j,E) = N_{ph}G^{<}_{j,j}(E-\hbar\omega_{ph}) + (N_{ph}+1)\,G^{<}_{j,j}(E+\hbar\omega_{ph}), \quad (19)$$

$$\overline{G}^{R}(x_j,E) = N_{ph}G^{R}_{j,j}(E-\hbar\omega_{ph}) + (N_{ph}+1)\,G^{R}_{j,j}(E+\hbar\omega_{ph}) + \tfrac{1}{2}\left[G^{<}_{j,j}(E-\hbar\omega_{ph}) - G^{<}_{j,j}(E+\hbar\omega_{ph})\right], \quad (20)$$

are the lesser and retarded Green's functions for the hole-optical phonon interactions, respectively. The term N_{ph} is the Bose–Einstein distribution for the phonon energy $\hbar\omega_{ph}$. During numerical evaluation of equations (17) to (20), the principal value of G^R and the real part of the self-energies are generally neglected. It was shown that this approximation introduced no significant error in the transport properties [20]. The charge and current criteria of convergence were set to 1%.

3.2. Evaluation of coefficients $F^{n,p}_{n',p'}$

In the hole-phonon interaction self-energy introduced in equation (17), the coefficients F carry information about the type of hole-phonon coupling and form factors due to the basis transformations, expressed as,

$$F^{n,p}_{n',p'} = \sum_d \int \mathrm{d}\mathbf{r}_\perp \zeta^\dagger_n(\mathbf{r}_\perp) D_d \zeta_p(\mathbf{r}_\perp)\cdot\zeta_{p'}(\mathbf{r}_\perp) D^\dagger_d \zeta_{n'}(\mathbf{r}_\perp), \quad (21)$$

where the x_j was omitted for brevity and $d = x, y, z$. Being a vector of size $\mathrm{N}_b \times 1$, $\zeta^\dagger_n$ is the real-space wave function for the mode n at x_j and . It is retrieved from the following basis transformation,

$$\zeta_n(\mathbf{r}_\perp) = U_K(\mathbf{r}_\perp)\Lambda_n, \quad (22)$$

where $\Lambda_n = U_M(n)$ is the mode-space wave function for the mode n. Finally, the integral in equation (21) is carried out over the cross/confinement section/direction of the nanostructure. Furthermore, in this work, D_d describes the hole-(acoustic/optical)-phonon interactions within the six-band $k\cdot p$ framework, being a 6×6 matrix. Herein, they are as given in [13], following the valence band deformation potential theory [22].

4. Simulations

The aforementioned mode-space full-band quantum transport approach has been implemented into the multi-scale transport tool called NESS [21]. We have efficiently simulated coherent and dissipative transport through a device consisting of two GAA in series and a p-type Si nanowire channel with a single dopant within each gated region. Both gates' work functions are 4.55 eV. The external source/drain bias is set to -0.1 V during this whole study. Figure 1 provides a detailed description of the device. The Si nanowire is 3 nm thick covered by 1 nm layer of SiO_2. The gate lengths are $L_{g1} = L_{g2} = 7$ nm, being 1 nm distance from each other. The source/drain is highly doped with $N_A = 10^{20}$ cm^{-3} and the region under both gates is considered intrinsic. The Coulomb potential of the dopants is described by adding a negative electron charge on the impurity site in the Poisson's equation. This approach correctly describes the long-range Coulombic tail, which is known to have the largest impact on the transport [23]. Electron-electron interaction is then only treated within the Hartree approximation by solving the NEGF transport equations self-consistently with Poisson's equation.

The source and drain Fermi energies are fixed at $E_{FS} = 0$ and $E_{FD} = 0.1$ eV, respectively. Then, the energy window considered for quantum transport simulations is chosen as $E_{FS} - 1.0\ \mathrm{eV} < E < E_{FD} + 0.2$ eV in order to ensure a good convergence. Luttinger parameters used for the six-band $k\cdot p$ Hamiltonian were taken from [15], where they were calibrated to sp^3 s* TB electronic calculations performed with the TCAD tool OMEN [19].

Finally, it is worth mentioning that the performance of the mode-space $k\cdot p$ approach presented in previous sections is limited by the number of modes N_M, size of the nanowire, and by how dense the energy grid is for the NEGF recursive algorithm. For instance, thicker nanowires would require us to increase N_M and to consider a denser energy grid. However, in comparison to real-space $k\cdot p$, the mode-space approach would still offer higher performance.

4.1. Band structure calculations

The K-space Hamiltonian has been discretized using a uniform mesh in all directions with mesh step $a_0 = 0.2$ nm, leading to $\tilde{N}_K = 225$. By applying the **QR** factorization approach, as explained in section 2, the size of the original problem is reduced, for instance, by approximately 60% in the case of a window energy of $E_v - E_n = 0.6$ eV, finding that only $N_K = 92$ modes are needed to fully reproduce the band structure with very good accuracy. The latter can be observed in figure 2, where the comparison of the full band structures of a 3 nm body thickness Si nanowire in the $\langle 111\rangle$ crystallographic direction was computed by means of the K-space and mode-space Hamiltonians. Note that for energies below $E \approx -1.7$ eV the mode-space approach becomes less accurate, particularly around $k_x = 0$. Regarding the transport problem in section 3, we have considered $N_M = 12\,(\ll N_K)$ modes after

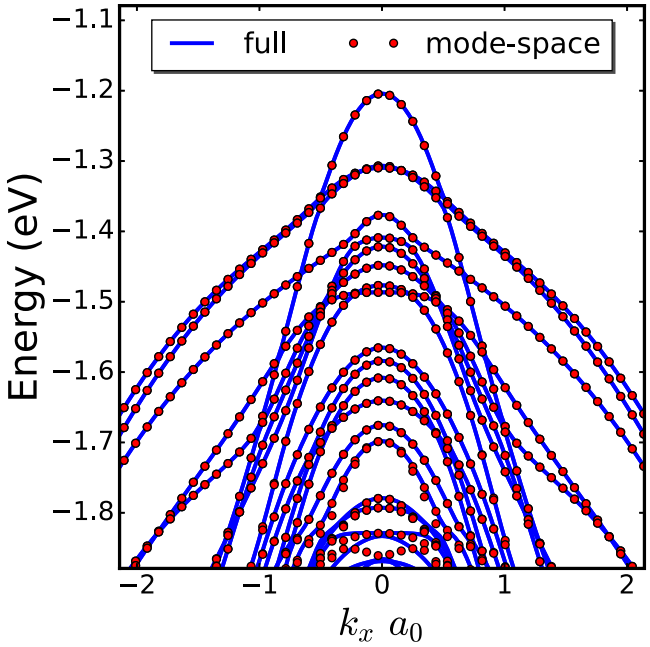

Figure 2. Full-band structures of a 3 nm body thickness Si nanowire in the ⟨111⟩ crystallographic direction. Band structures were calculated by means of the K-space and mode-space Hamiltonians. The latter was extracted within the window energy $E_v - E_n = 0.6$ eV. Note how the full K-space band structure can be recovered with very good accuracy above $E \approx -1.7$ eV. Valence band edge is set to $E_v = -1.12$ eV.

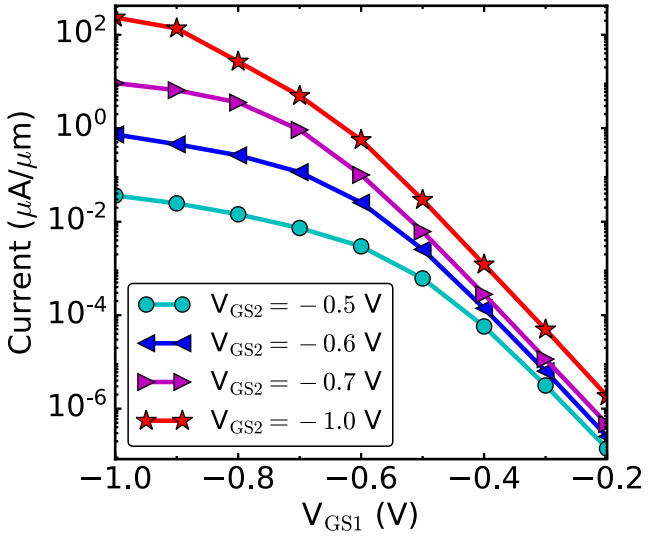

Figure 3. $I_D - V_{GS1}$ characteristics for different V_{GS2} applied at gate two. Source-to-drain applied bias is -0.1 V.

directly diagonalizing equation (8). Although including additional modes could lead to more accurate results, no substantial improvement of the accuracy was obtained. In particular, the latter number of modes was a good trade-off between accuracy and computational cost.

4.2. Ballistic limit

Figure 3 shows $I_D - V_{GS1}$ characteristics for different applied voltage at gate two, i.e. $V_{GS2} = -0.5$, -0.6, $-0.7$and -1.0 V. The lower the V_{GS2}, the higher the current in the whole range of V_{GS1} shown here. One can also observe that all curves start saturating at $V_{GS1} \lesssim -0.6$ V, whereas in the ON-state, the current increases approximately one-order of magnitude at each value of V_{GS2}. By applying a voltage at each gate, the effective potential barrier in the two-gated region is lowered and the dopant energy levels (DELs) are pushed up to higher energies. The latter are observed in the local density of states (LDOS) plotted in figure 4 for different states of the device in figure 1. One can observe how the DELs can be controlled by the applied gate voltages. The white line is the potential averaged with the hole concentration: $V_{av}(x) = \frac{1}{p_K(x)} \int d\mathbf{r}_\perp p(x, \mathbf{r}_\perp) V(x, \mathbf{r}_\perp)$.

Figure 4(a) shows the LDOS when the applied voltage is the same on both gates, $V_{GS1} = V_{GS2} = -0.6$ V. One can observe that the DELs of the second dopant are slightly shifted upwards by approximately 17 meV. The latter is due to the effect of the external source-to-drain bias. In equilibrium conditions, i.e. when $V_{DS} = 0$ V (not shown here), DELs from both dopants perfectly align to each other in terms of energy.

In figure 4(b), where $V_{GS1} = -0.6$ V and $V_{GS2} = -1.0$ V, the highest DELs of each dopant, are separated by approximately -0.18 eV despite the absolute value of V_{GS2} being increased by 0.4 V. This counter-intuitive phenomenon can be explained by the different screening mechanisms of each dopant. The dopant on the left is closer to the source, getting screened faster than the dopant on the right, which is mostly populated by the electrons injected from the drain. First, the decrease of the gate voltage increases the initial DEL, which is progressively populated by source/drain holes. This screening mechanism produces an opposite effect of level lowering, and the DEL slowly saturates towards the average potential in the source/drain, as observed in figure 5. When the gate voltage is strong enough, the DEL can go above the average potential in the source/drain. This is the case of $V_{GS1} = -0.9$ V and $V_{GS2} = -1.0$ V, in figure 5, where one can see how the DEL of the first dopant abruptly goes above the average potential in the source. Meanwhile, the DEL of the second dopant slightly changes over the V_{GS1} range for all values of V_{GS2} considered in this work.

4.3. Impact of hole-phonon interactions

This section reports on the impact of hole-phonon interactions on the ballistic simulation results presented above. The quantum transport simulations, in the presence of phonon scattering, were carried out by means of the NEGF and SCBA formalisms within a mode-space representation of the six-band $k \cdot k$ Hamiltonian introduced in section 3. Simulations are presented below.

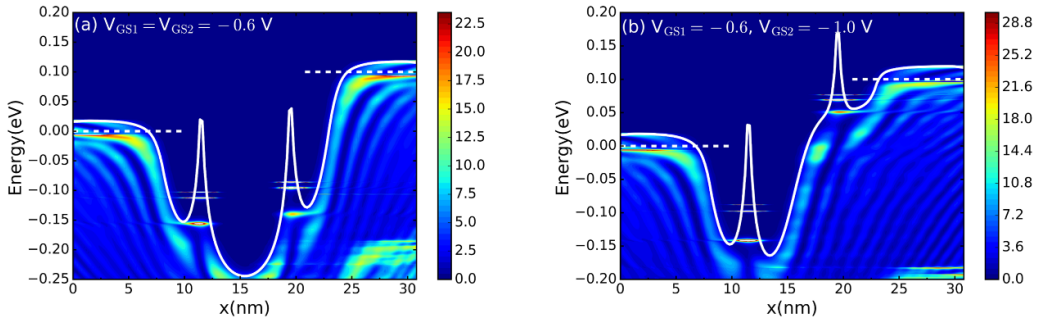

Figure 4. LDOS for two configurations of V_{GS1} and V_{GS2} in figure 3. For comparison purposes, V_{GS1} is fixed at -0.6 V whereas V_{GS2} is (a) -0.6 V and (b) -1.0 V. LDOS in units of eV^{-1}. White line and horizontal dashed-lines are the potential averaged with the hole concentration and source/drain Fermi energies, respectively.

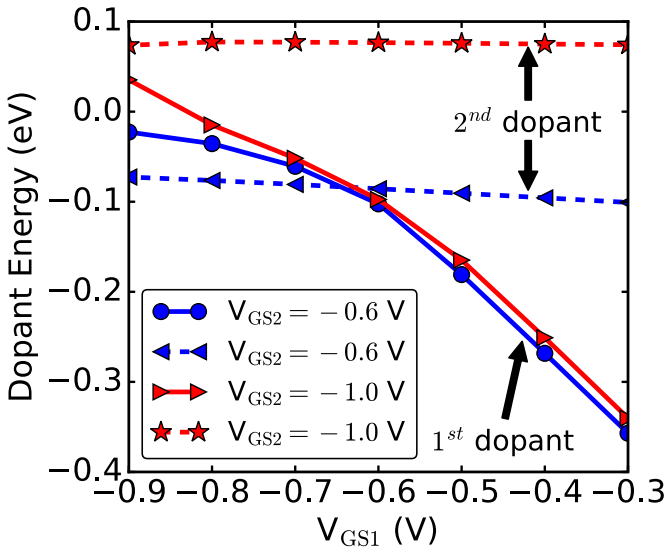

Figure 5. Dopant energy levels as a function of the V_{GS1} and different V_{GS2} with respect to the equilibrium Fermi energy, i.e. $E_F = 0$ eV.

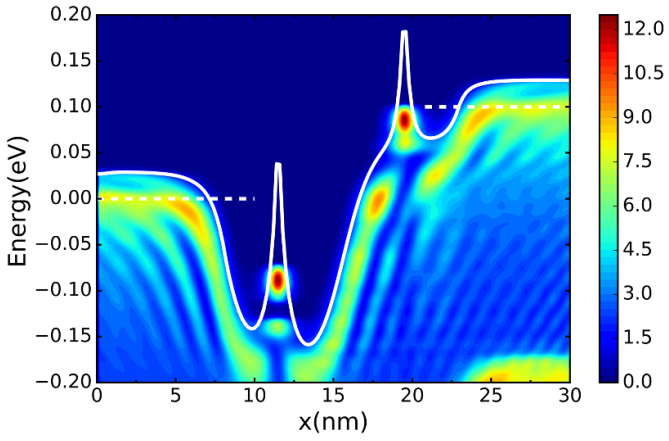

Figure 6. LDOS in the presence of hole-phonon interactions at $V_{GS1} = -0.6$ V and $V_{GS2} = -1.0$ V. Units eV^{-1}. White line and horizontal dashed-lines are the potential averaged with the hole concentration and source/drain Fermi energies, respectively.

Figure 6 shows the LDOS in the presence of hole-phonon interactions for the same gate voltage configuration as its ballistic counterpart in figure 4(b) at $V_{GS1} = -0.6$ V and $V_{GS2} = -1.0$ V. Due to phonon scattering, the LDOS is globally broadened, as observed around the DELs. Note that in the ballistic regime (figure 4(b)), three sharp DELs exist. However, by switching the phonon scattering on, the highest two DELs are merged in one broader energy level. One can also observe that there is an overlapping of energy levels of the second dopant.

A stronger effect of phonon scattering can be seen in the current spectrum, as reported in figure 7, which compares spectra corresponding to the ballistic regime (figure 7(a)) and the one in the presence of hole-phonon interactions (figure 7(b)) for $V_{GS1} = -0.6$ V and $V_{GS2} = -1.0$ V. In the case of the ballistic, we can observe three resonant current branches within the window energy of -80 and -150 meV. Each branch corresponds to the energy levels of the first dopant, as shown in the LDOS plotted on figure 4(b). The fourth and broadest branch in figure 7(a) below the potential barrier carries the current contribution of thermionic holes injected from the source and drain regions. In the presence of phonon scattering, holes injected from the source can relax into the resonant energy levels by emitting a single or multiple phonons. The latter is observed in figure 7(b), where the resonant sites help holes to relax to higher energies, starting from the first dopant site up to the drain region. Interestingly, holes that are backscattered can also relax to higher energies in the direction of the source region.

Figure 8 shows the impact of hole-phonon interactions on the $I_D - V_{GS1}$ characteristics for two different configurations: $V_{GS2} = -0.6$ V (blue lines) and $V_{GS2} = -1.0$ V (red lines). When compared to the corresponding ballistic case, the current is reduced over the entire range of V_{GS1} shown here. Their losses are plotted in the inset figure with the same color code and V_{GS1} range. The percentage loss is defined as $|I_{ball} - I_{scatt}|/I_{ball} \times 100\%$. Similar to current p-type MOSFETs, the lower the V_{GS1}, the stronger the impact of phonon scattering causing a reduction in the current and therefore enhancing the losses. For the transistor considered in this work, this trend only happens for high voltages, i.e $V_{GS1} \geqslant -0.7$ V, within the sub-threshold part of the $I_D - V_{GS1}$ characteristics. The loss

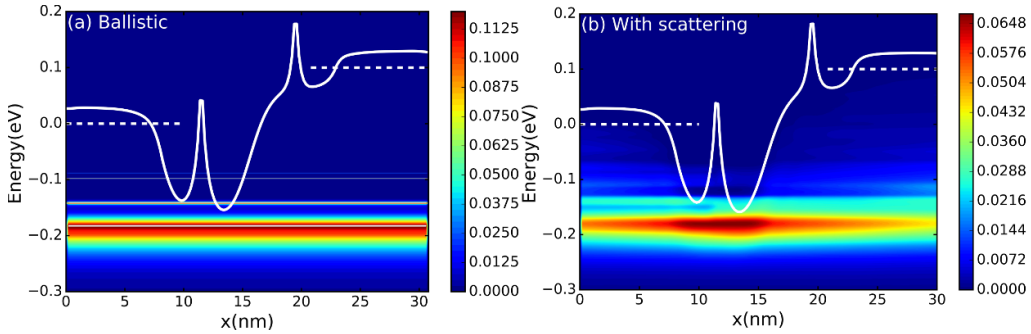

Figure 7. Comparison of current spectra within (a) the ballistic regime and (b) in the presence of the hole-phonon interactions at $V_{GS1} = -0.6$ V and $V_{GS2} = -1.0$ V. Current spectra in units of $\mu A/eV$. White line and horizontal dashed lines are the potential averaged with the hole concentration and source/drain Fermi energies, respectively.

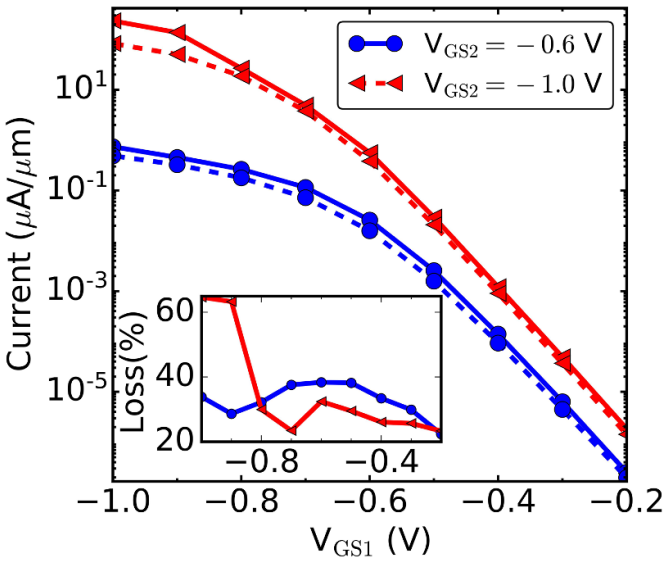

Figure 8. Comparison of the ballistic (solid lines) and hole-phonon interaction (dashed lines) $I_D - V_{GS1}$ characteristics at $V_{GS2} = -0.6$ V (blue lines) and $V_{GS2} = -1.0$ V (red lines). Source-to-drain applied bias is -0.1 V. Inset figure shows their respective current loss defined as $|I_{ball} - I_{scatt}| \times 100\%/I_{ball}$.

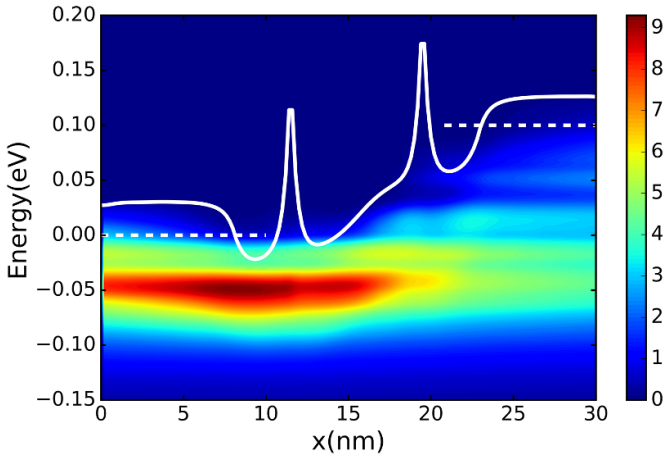

Figure 9. Current spectrum in the presence of the hole-phonon interactions at $V_{GS1} = -1.0$ and $V_{GS2} = -0.6$ V. Units $\mu A/eV$. White line and horizontal dashed-lines are the potential averaged with the hole concentration and source/drain Fermi energies, respectively.

has a slight decrement at $V_{GS1} = -0.9$ V and $V_{GS1} = -0.7$ V in the case of $V_{GS2} = -0.6$ V and $V_{GS2} = -1.0$ V, respectively. The latter could be explained by a substantial contribution of the tunneling current flowing through the resonant site of the first dopant when its energy level gets closer to the Fermi level. For lower V_{GS1}, optical-phonon scattering is expected to become dominant, contributing to an increase in the loss, which reaches up to 60% in the case of $V_{GS1} = -1.0$ V and $V_{GS2} = -1.0$ V.

Figure 9 shows the current spectrum for the previous configuration where the strong effect of the phonon scattering can be seen. In this case, the DEL is below the source band edge, creating a resonant site where holes injected from the source can tunnel to the drain. A second current branch around the energy level of the first dopant is then generated. Due to the strong electric field within the gate regions, it quickly fades due to the large number of holes relaxing to higher energies until they reach the top of the drain band edge.

5. Conclusion

A mode-space-based full-band quantum transport simulator with phonon scattering using the six-band $k \cdot p$ method has been developed and implemented in Nano-Electronic Simulation Software (NESS) from the University of Glasgow. Based on the non-equilibrium Green's function formalism and self-consistent Born's approximation, an expression for the hole-phonon interaction self-energy within the mode-space representation was introduced. We efficiently simulated a transistor-like device, which may be relevant for future quantum technologies, formed by two GAA in series and a p-type Si nanowire channel with a single dopant within each gated region. The dopant's Coulomb potentials were mimicked by adding a negative electron charge on the impurity site in the 3D Poisson's equation.

Our findings suggest that hole-phonon interactions have a significant detrimental impact on the transistor performance when compared to ballistic simulations. Here, we have reported that the current could be reduced by up to 60%, depending on the gate voltage configuration.

Acknowledgments

The authors are most grateful to Mathieu Luisier for granting us access to the IIS computer infrastructure to perform the simulations leading to the results presented in this work. This research was funded by the Engineering and Physical Sciences Research Council (EPSRC), through-Grant Nos. EP/S001131/1 and EP/P009972/1. This project has also received funding from the EPSRC Impact Acceleration Account scheme under Grant Agreement No. EP/R511705/1 (Nano-Electronic Simulation Software (NESS)—creating the first open source TCAD platform in the world and Fast Track - Development boost for the Device Modelling group open-source NESS computational framework).

Author contributions

Conceptualization, HCN; methodology, HCN; software writing and implementation, HCN and CMB; original draft preparation, HCN; writing review and editing, HCN; CMB - writing and editing; supervision, VPG and AA; funding acquisition, CMB, VPG and AA. All authors have read and agreed to the published version of the manuscript.

Appendix A. $k\cdot p$ matrix coefficients for different crystallographic directions

In section 2, the six-band $k\cdot p$ Hamiltonian hasbeen given in terms of the three-band $k\cdot p$ Hamiltonian in $\{|X\rangle,|Y\rangle,|Z\rangle\}$ basis. As expressed in equation (1), information about the crystallographic direction is contained in the following group of matrices:

$$A=\begin{pmatrix} c_t^2c_p^2 & 0 & 0\\ 0 & c_t^2s_p^2 & 0\\ 0 & 0 & s_t^2\end{pmatrix},\ B=\begin{pmatrix} s_p^2 & 0 & 0\\ 0 & c_p^2 & 0\\ 0 & 0 & 0\end{pmatrix}, \tag{23}$$

$$C=\begin{pmatrix} s_t^2c_p^2 & 0 & 0\\ 0 & s_t^2s_p^2 & 0\\ 0 & 0 & c_t^2\end{pmatrix},\ D=2c_tc_ps_p\begin{pmatrix} 1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}, \tag{24}$$

$$F=2s_pc_ps_t\begin{pmatrix} -1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},\ G=-2c_ts_t\begin{pmatrix} c_p^2 & 0 & 0\\ 0 & s_p^2 & 0\\ 0 & 0 & -1\end{pmatrix}, \tag{25}$$

$$A'=\begin{pmatrix} 0 & c_t^2c_ps_p & -c_tc_ps_t\\ '' & 0 & c_ts_ts_p\\ '' & '' & 0\end{pmatrix},\ B'=\begin{pmatrix} 0 & -c_ps_p & 0\\ '' & 0 & 0\\ 0 & 0 & 0\end{pmatrix}, \tag{26}$$

$$C'=\begin{pmatrix} 0 & s_t^2c_ps_p & c_ts_tc_p\\ '' & 0 & -c_ts_ts_p\\ '' & '' & 0\end{pmatrix},$$

$$D'=\begin{pmatrix} 0 & c_t(s_p^2-c_p^2) & -s_ts_p\\ '' & 0 & -c_ps_t\\ '' & '' & 0\end{pmatrix}, \tag{27}$$

$$F'=\begin{pmatrix} 0 & s_t(c_p^2-s_p^2) & -s_pc_t\\ '' & 0 & -c_tc_p\\ '' & '' & 0\end{pmatrix},$$

$$G'=\begin{pmatrix} 0 & -2c_tc_ps_ts_p & c_p(s_t^2-c_t^2)\\ '' & 0 & s_p(c_t^2-s_t^2)\\ '' & '' & 0\end{pmatrix}, \tag{28}$$

where s_α and c_α represent $\sin\alpha$ and $\cos\alpha$, respectively. For the sake of brevity, symmetric matrix elements have been omitted (''). In the spirit of a spherical coordinate system (k,θ,ϕ), ϕ $(\alpha=p)$ is the azimuthal angle between k_x-axis and the projection of the $k-$ vector onto the k_x-k_y plane. Whereas, θ $(\alpha=t)$ is the angle formed by the latter and the direction to which the $k-$ vector points. Both angles can be then extracted from,

$$\cos\phi=\frac{1}{l_x}\frac{1}{d_{l_xl_y\infty}},\quad \cos\theta=\frac{d_{l_xl_y\infty}}{d_{l_xl_yl_z}}, \tag{29}$$

where l_i is the corresponding Miller index for $i=x,y,z$ and the reciprocal distance $d_{l_xl_yl_z}$ has been defined as,

$$d_{l_xl_yl_z}=\sqrt{\sum_{i=x,y,z}l_i^{-2}}. \tag{30}$$

For the particular case in this work, $l_x=1$, $l_y=1$ and $l_z=1$, one can find that $c_p=1/\sqrt{2}$ and $c_t=\sqrt{2/3}$. Finally, by means of the set of equations (23)–(30), other relevant crystallographic directions can be then easily incorporated in equation (1) for quantum transport simulations.

Appendix B. Discretization of the $k\cdot p$ Hamiltonian in K-space

The real-space Hamiltonian given by equations (1)–(5) is first unitary transformed to K-space via [15],

$$H_{K,L,L'}(x)=\int \mathrm{d}\Omega\, U_{K_L}(\mathbf{r}_\perp)H_R(-i\nabla,\mathbf{r})U_{K_{L'}}(\mathbf{r}_\perp), \tag{31}$$

where $\mathrm{d}\Omega=\mathrm{d}y\mathrm{d}z$ and

$$U_{K_L}(\mathbf{r}_\perp) = U(\mathbf{r}_\perp, \mathbf{k}_L) = \frac{2}{L_y L_z}\sin(k_p y)\sin(k_q z). \quad (32)$$

$\mathbf{k}_L = \pi(p/L_y, q/L_z)$ is evaluated in the k-space grid point $L = (p, q)$. L_y and L_z are the cross-section lengths in y- and z-directions, with real-space coordinates $\mathbf{r}_\perp = (y, z)$, respectively. Carrying out the integral over the cross-sectional plane in equation (31), one straightforwardly obtains,

$$K^2_{y(z)} = \int \mathrm{d}\Omega\, U_{K_L} k^2_{y(z)} U_{K_{L'}} = k^2_{p(q)}\delta_{pp'(qq')}, \quad (33)$$

$$K_y = \int \mathrm{d}\Omega\, U_{K_L} k_y U_{K_{L'}} = -\tfrac{4i}{\pi} k_{p'} \tfrac{p}{p^2 - p'^2}\delta_{qq'}, \quad (34)$$

$$K_z = \int \mathrm{d}\Omega\, U_{K_L} k_z U_{K_{L'}} = -\tfrac{4i}{\pi} k_{q'} \tfrac{q}{q^2 - q'^2}\delta_{pp'}. \quad (35)$$

After a proper discretization in the transport x-direction, with N_x cross-sectional planes, the K-space Hamiltonian can be written as,

$$H_K = \begin{pmatrix} H'_{0\,\mathrm{K}} & 0 \\ 0 & H'_{0\,\mathrm{K}} \end{pmatrix} + \begin{pmatrix} W'_K + W'^{\dagger}_K & 0 \\ 0 & W'_K + W'^{\dagger}_K \end{pmatrix} + \begin{pmatrix} \mathbf{M} & \mathbf{N} \\ -\mathbf{N}^* & \mathbf{M}^* \end{pmatrix}\delta_{pp'}\delta_{qq'} + \begin{pmatrix} V'_K & 0 \\ 0 & V'_K \end{pmatrix}, \quad (36)$$

which briefly can be re-expressed as follows,

$$H_K(j) = H_{0\,\mathrm{K}}(j) + W_K(j) + W^{\dagger}_K(j) + V_K(j), \quad (37)$$

for the j-th cross-sectional plane. In the latter, the spin–orbit coupling interactions have been included in its first term on the right-hand side. Ultimately, in equation (36),

$$\begin{aligned} H'_{0\,\mathrm{K}} =& M\left(\frac{2}{\Delta x^2} + k_p^2 + k_q^2\right)\delta_{pp'}\delta_{qq'}\mathbf{I}_{3\times 3} \\ &+ (L - M)\left(A\frac{2}{\Delta x^2}\delta_{pp'}\delta_{qq'} + BK_y^2 + CK_z^2 + FK_yK_z\right) \\ &- N\left(A'\frac{2}{\Delta x^2}\delta_{pp'}\delta_{qq'} + B'K_y^2 + C'K_z^2 + F'K_yK_z\right), \end{aligned} \quad (38)$$

$$\begin{aligned} W'_K =& -\frac{M}{\Delta x^2}\delta_{pp'}\delta_{qq'}\mathbf{I}_{3\times 3} \\ &+ (L - M)\left(A\frac{-1}{\Delta x^2}\delta_{pp'}\delta_{qq'} + D\frac{-i}{2\Delta x}K_y + G\frac{-i}{2\Delta x}K_z\right) \\ &- N\left(A'\frac{-1}{\Delta x^2}\delta_{pp'}\delta_{qq'} + D'\frac{-i}{2\Delta x}K_y + G'\frac{-i}{2\Delta x}K_z\right), \end{aligned} \quad (39)$$

and the potential V'_K in K-space can be numerically evaluated, for instance, as,

$$V_{K,L,L'}(j) = \sum_{i=1}^{N_y N_z} \Delta\Omega_i U_{K_L}(x_j, \mathbf{r}_i) V(x_j, \mathbf{r}_i) U_{K_{L'}}(x_j, \mathbf{r}_i), \quad (40)$$

where $\mathbf{r}_i = (y_m, z_n)$ is the i-th grid point in real space.

ORCID iD

Hamilton Carrillo-Nuñez
https://orcid.org/0000-0002-5042-3664

References

[1] Tomioka K, Yoshimura M and Fukui T 2012 A iii–v nanowire channel on silicon for high-performance vertical transistors *Nature* **488** 189–92
[2] Robertson J and Wallace R M 2015 High-k materials and metal gates for CMOS applications *Mater. Sci. Eng. R: Rep.* **88** 1–41
[3] Ritzenthaler R *et al* 2018 Vertically stacked gate-all-around Si nanowire CMOS transistors with reduced vertical nanowires separation, new work function metal gate solutions and dc/ac performance optimization *2018 IEEE Int. Electron Devices Meeting (IEDM)* p 21.5.1–21.5.4
[4] Int. Roadmap for Devices and Systems *More Moore* 2017
[5] Maurand R *et al* 2016 A CMOS silicon spin qubit *Nat. Commun.* **7** 13575
[6] Roche B, Dupont-Ferrier E, Voisin B, Cobian M, Jehl X, Wacquez R, Vinet M, Niquet Y-M and Sanquer M 2012 Detection of a large valley-orbit splitting in silicon with two-donor spectroscopy *Phys. Rev. Lett.* **108** 206812
[7] Niquet Y M, Lherbier A, Quang N H, Fernández-Serra M V, Blase X and Delerue C 2006 Electronic structure of semiconductor nanowires *Phys. Rev.* B **73** 165319
[8] Luisier M, Schenk A, Fichtner W and Klimeck G 2006 Atomistic simulation of nanowires in the $sp^3d^5s^*$ tight-binding formalism: From boundary conditions to strain calculations *Phys. Rev.* B **74** 205323
[9] Hohenberg P and Kohn W 1964 Inhomogeneous electron gas *Phys. Rev.* **136** B864–B871
[10] Kohn W and Sham L J 1965 Self-consistent equations including exchange and correlation effects *Phys. Rev.* **140** A1133–A1138
[11] Yu P and Cardona M 2005 *Fundamentals Semiconductors: Phys. Mater. Properties* (Berlin: Springer) p 68 3rd edn
[12] Kane E O 1956 Energy band structure in p-type germanium and silicon *J. Phys. Chem. Solids* **1** 82–99
[13] Cavassilas N, Michelini F and Bescond M 2011 Multiband quantum transport simulations of ultimate p-type double-gate transistors: Effects of hole-phonon scattering *J. Appl. Phys.* **109** 073706
[14] Pons N, Cavassilas N, Raymond L, Michelini F, Lannoo M and Bescond M 2011 Three-dimensional $k \cdot p$ real-space quantum transport simulations of p-type nanowire transistors: Influence of ionized impurities *Appl. Phys. Lett.* **99** 082113
[15] Shin M 2009 Full-quantum simulation of hole transport and band-to-band tunneling in nanowires using the $k \cdot p$ method *J. Appl. Phys.* **106** 054505
[16] Huang J Z, Chew W C, Peng J, Yam C, Jiang L J and Chen G 2013 Model order reduction for multiband quantum transport simulations and its application to p-type junctionless transistors *IEEE Trans. Electron Devices* **60** 2111–19
[17] Shin M, Lee S and Klimeck G 2010 Computational study on the performance of Si nanowire P-MOSFETs based

on the $k \cdot p$ method *IEEE Trans. Electron Devices* **57** 2274–83

[18] Ziegler A and Luisier M 2017 Phonon confinement effects in diffusive quantum transport simulations with the effective mass approximation and $k\cdot$ method *2017 Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD)* pp 25–8

[19] Luisier M and Klimeck G 2009 Atomistic full-band simulations of silicon nanowire transistors: Effects of electron-phonon scattering *Phys. Rev.* B **80** 155430

[20] Svizhenko A and Anantram M P 2003 Role of scattering in nanotransistors *IEEE Trans. Electron Devices* **50** 1459–66

[21] Berrada S, Carrillo-Nuñez H, Dutta T, Duan M, Adamu-Lema F, Lee J, Georgiev V, Medina-Bailon C and Asenov A 2018 Ness: new flexible nano-electronic simulation software *2018 Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD)* pp 22–5

[22] Hinckley J M and Singh J 1990 Hole transport theory in pseudomorphic $si_{1-x}ge_x$ alloys grown on Si(001) substrates *Phys. Rev.* B **41** 2912–26

[23] Bescond M, Lannoo M, Raymond L and Michelini F 2010 Single donor induced negative differential resistance in silicon n-type nanowire metal-oxide-semiconductor transistors *J. Appl. Phys.* **107** 093703